

特集：臨床研究実施の現状と課題

第2部 医療データベースの疫学研究への活用：医療データベースの研究利用と社会還元

DPC データに基づく臨床研究

東京大学大学院医学系研究科 医療経営政策学

康 永 秀 生

はじめに

患者集団におけるなんらかのアウトカム評価を行う際、アウトカムに影響する多様な因子を統計的に調整するためには、一定以上の症例数が必要となる。医療技術の進歩のスピードはしだいに速まっており、短期間で一定の症例数を集め評価する必要性がますます高まっている。そのために臨床疫学研究の多施設化が世界的に進行している。

言うまでもなく、臨床疫学研究の王道はランダム化比較試験 (randomized controlled trial : RCT) である。薬剤の治験や新治療技術の効果判定などにおいて、RCT の果たす役割は大きい。しかし、RCT は、①ばく大なコストがかかる、②研究テーマが限定される、とりわけすでに普及している医療技術には適用されにくい、③倫理的制約がある、といった問題を抱えており、実際の適用範囲は限られている。また、RCT においては、exclusion criteria を設けて選抜されたサンプルが実臨床における患者母集団とはかけ離れてしまう、という批判もある。

エビデンスは日常臨床のなかに潜んでいる。現場で時々刻々と発生する臨床データを広汎かつ効率的に拾い上げ、臨床疫学研究に活かすというアプローチは今後ますます重要となるだろう。そのひとつの方法論として、疾患別の登録型データベース (registry database) があげられる。学会などアカデミアが主体となり、所属医療機関・医師の協力を得て、共通のデータ・フォーマットを用いて、特定の疾患や診療領域の患者個票データを多施設から恒常的に登録してもらう方式である。もうひとつの方法論が、

医療機関の日常業務として発生する診療報酬請求データ (administrative claim data) を臨床研究への利用に供するというアプローチである。

本稿では、日本版診断群分類 (Diagnostic Procedure Combination : DPC) に基づく診療報酬請求データベースについて概説し、それを用いた臨床研究の有用性と課題について論じる。

1 DPC データベースを用いた臨床研究

米国では 1990 年代から administrative claim database が整備されてきた。Nationwide Inpatient Sample (NIS) は、2007 年時点で全米 40 州 1044 病院が参加し、全米の入院患者の約 20% (年間約 800 万人) を占める巨大なデータベースである。Medicare Claim Database は、米国の 65 歳以上を対象とした公的医療保険 Medicare に加入している全入院患者の診療報酬データベースである。これらのデータベースを用いた臨床研究の例は枚挙にいとまがない。

日本では、2002 年から DPC データ調査研究班 (<http://www.dpcsg.jp/>) が主体となって、DPC データの収集・分析を進めてきた。調査参加施設は年々増加しており、2009 年度には 818 施設、のべ退院患者数 257 万人に達した。この数は、同時期の日本全国の一般病床退院患者の約 40% を占める (表 1)。

DPC データベースの「様式 1」には、医療機関情報 (施設コード、開設主体、病床規模)、患者基本情報 (年齢、性別、入退院日、入退院経路、予定・緊急入院の別、救急搬送の有無、在院日数、退院時転帰など)、診断情報 (診断名、入院時併存症、入院後

表 1 DPC データ調査研究班・参加医療施設

年度	調査期間	参加医療施設数	のべ退院患者数 (万人)
2002	7～10 月	82	26
2003	7～10 月	185	44
2004	7～10 月	174	45
2005	7～10 月	249	73
2006	7～12 月	262	108
2007	7～12 月	926	299
2008	7～12 月	855	286
2009	7～12 月	818	257

合併症), 手術情報 (手術日, 術式, 麻酔法) などに加えて, 広範な診療関連情報 (各種疾患の重症度分類など) も含まれる。「EF ファイル」は診療行為明細であり, 個々の患者に入院中に実施された検査・処置, 使用された薬剤・特定保険医療材料の種類・量・日付が網羅されている。麻酔時間, 輸血量, 医療費などのデータも含まれる。

DPC データベースが単独で提供できるデータは記述疫学ならびに後ろ向きの観察疫学データにすぎず, この点は RCT に及ぶところではない。また, 疾患登録データベースと比較すれば, 含まれるデータ項目の内容は限られる。DPC データベースは入院中のデータに限られるため, 術後合併症・在院死亡・在院日数などのアウトカムは得られるが, 長期予後に関するデータは得られない。診断や治療に関する詳細なプロセス・データは得られるが, 検査結果や治療効果に関するアウトカム・データは少ない。

しかし, DPC データベースの 2 つの利点は, 1 つは 800 以上の施設から収集された大規模データである点, もう 1 つはあらゆる疾患の入院患者に関するデータが含まれる点である。これまで日本では, 各種の疾病・治療に関する全国統計が不足していた。DPC データベースを用いれば, 共通フォーマットを用いた多施設の臨床データの集計が可能である。

これまで DPC データベースを用いた研究例として, ①各種疾患の記述疫学, ②治療後合併症や在院死亡率の統計, ③薬剤の有害事象の統計, ④薬剤の治療効果分析, ⑤外科手術の volume-outcome relationship 分析などがあげられる。報告された論文一覧は以下の URL を参照されたい (<http://plaza.umin.ac.jp/~hmp/cgi-bin/wiki/wiki.cgi>)。

2 DPC データベースの今後

DPC データベースは今後, 他のデータベースとリンクすることにより, さらにその有用性を高めていくことが期待される。

米国の Medicare Claim Database は, Surveillance, Epidemiology and End Results (SEER) program of cancer registries と呼ばれるがん登録データベースと共通の患者 ID でリンクされている。これにより, がんについては入院時の治療や短期のアウトカムのみならず, その後の生存率まで追跡が可能となっている (SEER-Medicare Linked Database)。DPC データベースと他のデータベースのリンクを実現するには, 共通の患者 ID を用いるための技術的・倫理的・財政的諸問題がクリアされなければならない。

日本でも今後, 学会レベルでの症例登録データベースの構築が進められていくことが期待されている。新規のデータベースと DPC データベースをリンクすることは, 次の点で有用である。すなわち, すでに各施設でルーチンワーク化している DPC の入力作業と, 新データベースの入力作業の重複を避け, 省力化による各施設の負担軽減が可能となる。DPC における病名情報やプロセス情報を共有化でき, なおかつ DPC データにはない詳細なアウトカム・データを付加することができる。個々のデータベースが単独では実現しえない利点を双方が享受できる。

おわりに

日本の医学研究は基礎医学に偏重し続けてきた。基礎研究は 1 施設でもできる。しかし, 臨床疫学研究は, 多施設からデータを収集する必要がある。臨床医だけで大規模データの収集はできない。疫学・公衆衛生学などの社会医学研究者が臨床医と連携して, データ・マネジメントや統計分析を実践しなければならない。しかし, 日本では伝統的に, 社会医学研究者と臨床医との交流が活発とはいえなかった。そのことが一因となり, 日本の臨床疫学研究が欧米の後塵を拝する結果となった, といっても過言ではなかろう。今後, 日本の臨床疫学研究が世界に伍して持続的に発展を続けていくためには, 臨床医学と社会医学の連携強化が不可欠である。DPC データベースは, 上記のような日本の臨床疫学研究の閉塞した状況を打開するひとつの突破口となりうるだろう。